EX PARTE OR LATE FILED

ORIGINAL

October 14, 1999

RECEIVED

OCT 14 1999

FEDERAL COMMUNICATIONS COMMISSION
OFFICE OF THE SECRETARY

Ms. Magalie Roman Salas
Secretary
Federal Communications Commission
445 Twelfth Street, S.W.
Washington, D.C. 20554

Re: CC Docket No. 96-45

Dear Ms. Salas,

This letter reports an ex parte contact in CC Docket 96-45. On October 14, 1999, I contacted Chuck Keller, Deputy Division Chief of the Accounting Policy Division, Common Carrier Bureau, and described a study entitled The Rural Differential: An Analysis of Population Demographics in Areas Served by Rural Telephone Companies ("Rural Study"). The Rural Study, attached to this letter, is hereby submitted for the record.

The Rural Study, authored by the Rural Policy Research Institute at the University of Missouri-Columbia and the Office of Social and Economic Data Analysis at the University of Missouri, contains data and findings relevant to the Commission's development of Universal Service Fund rules in this proceeding.

The Study analyzes data to show, among other things, what part of the U.S. population classified as "rural" under census data is actually served by non-rural telephone companies. Based upon information in the Rural Study, Table 1, about 20% of the U.S. population is "rural" by census definition and is served by non-rural telephone companies. Likewise, about 5% of the U.S. population is "rural" by census definition and is served by rural telephone companies. Rural Study, p. 2. This means that for every rural customer who is served by rural companies, there are four who are served by non-rural companies. The rural population of this country is served predominantly by non-rural telephone companies. In developing its Universal Service Fund rules, the FCC should strive to treat all rural customers fairly, regardless of whether they are served by a rural telephone company.

Sincerely,

Peter Bluhm/ehr

Peter Bluhm
Vermont Public Service Board
112 State Street, Drawer 20
Montpelier, Vermont 05620-2601
802/828-2358

Enclosure
G:\101084\1\RUPRIL~1.DOC

No. of Copies rec'd 0+4
List ABCDE

Rural Policy Research Institute
IOWA STATE UNIVERSITY • UNIVERSITY OF MISSOURI • UNIVERSITY OF NEBRASKA

The Rural Differential: An Analysis of Population Demographics in Areas Served by Rural Telephone Companies

August, 1999

P99-8

by

Vicki M. Hobbs
Chair, Telecommunications Panel
Rural Policy Research Institute (RUPRI)
University of Missouri-Columbia

and

John Blodgett
Office of Social and Economic Data Analysis
University of Missouri

For more information, contact:
RUPRI OFFICE
University of Missouri
200 Mumford Hall
Columbia, MO 65211
(573) 882-0316
FAX {573] 884=5310
office@rupri.org
www.rupri.org

The Rural Policy Research Institute provides objective analyses and facilitates dialogue concerning public policy impacts on rural people and places.

Absent from much of the discussion about rural America and the extent to which universal service programs and policies impact its residents is an understanding of the population served by rural telephone companies[1] and impacted by any change in universal service polices. It has been assumed that the demographics of those areas served by rural telcos were synonymous with the demographics of rural America. In reality, the 7% of the US population served by rural telcos may not necessarily typify the 61,656,000[2] or 25% of all persons classified as rural in the 1990 Census nor can any generalizations made be certain to apply to any particular geographic area. As important as the distinctions between those areas served by rural and non-rural telcos are the distinctions which can be drawn *among* those areas served by rural telcos.

As part of this analysis of the 'rural differential', telephone area code and exchange (NPA/NXX) were added to the search variables of the Basic Tables Generator maintained by the Office of Social and Economic Data Analysis at the University of Missouri. The development of this capability, funded by the Rural Policy Research Institute (RUPRI), in response to the anticipated needs of the Rural Task Force, enables web-based access[3] to 1990 census data (STF3 files) for any telephone exchange based on the census data of the wire center to which it is linked. (See pages 20-30 for an in-depth description of the methodology employed.) This effort was also undertaken as a precursor for the 2000 census data, which, when available, will yield parallel but more demographically current results. This empowering information can enable better understanding of local demographic circumstances as they relate to the provision of telecommunications services and of the potential local impact of changes in universal service support mechanisms. Certainly, reported "averages" or "means" do not do justice to the multiplicity of circumstances existing within the geographic regions of the U.S. served by rural telcos. Ultimately, it is only by understanding discreet service area distinctions that a full understanding of the mosaic of rural America can be achieved with respect to the provision of telecommunications.

Several important demographic characteristics highlight the significant differences between those areas served by rural and non-rural telcos. For illustrative purposes, several differences *within* telco service areas are also examined at the state level. No attempt is made herein to report all available demographic data by telco service area or to include all data accessible by wire center, state or region. Full, user-definable access to the complete dataset is available through the RUPRI website http://www.rupri.org.

---

[1] Rural telcos is this context refers to "Rural Telephone Companies" which incorporates the statutory definition in Section 3 (47) of the Telecommunications Act of 1996 (47 U.S.C. § 153 (37) and its application in the FCC rules, adopted pursuant to CC 96-45, which set a separate schedule and additional scrutiny for "rural telephone companies" than for "non-rural telephone companies". The FCC has recognized the self-certified rural telephone companies listed in the FCC Public Notice of December 31, 1998, DA 98-2642, as Eligible Telecommunications Carriers (ETCs).

[2] Statistical Abstract of the U.S., 1996

[3] The URL for accessing the census database by telephone exchange is http://www.rupri.org. Select "Demographic Data by Telco Exchange". [Do note that, while the data can be accessed by NPA/NXX (Area Code/Phone Exchange), the data reflect that of the wire center which may, when noted, include multiple exchanges.]

Population Density

Rural telcos serve approximately 7% of the population in the U.S., which at the time of the 1990 census, accounted for more than 17 million people. Of those 17 million, 73% are considered rural[4] by U.S. census definition, that is, residing in open country or in places of less than 2500 population. Of those 73% rural, however, more than nine out of ten (91%) live in a non-farm setting, refuting the still prevalent conventional wisdom that rural America is largely agricultural.

Just as rural telcos as defined by federal law do not exclusively serve "rural areas", neither do non-rural telcos exclusively serve urban or metropolitan areas. Twenty-one percent of the population served by non-rural telcos is classified as "rural" by census definition.

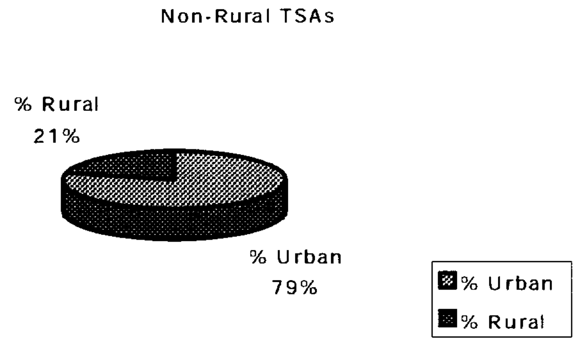
Table 1: Urban and Rural Composition of Rural and Non-Rural Telephone Service Areas

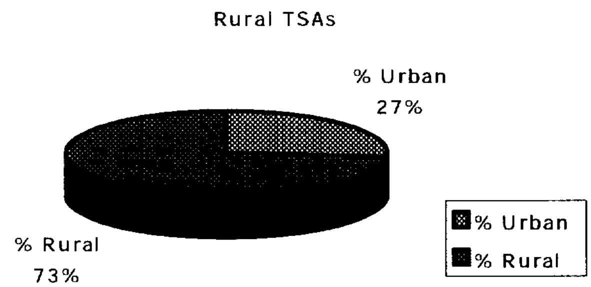
	Rural Service Areas	No-Rural Service Areas
% of U.S. Population	7%	93%
% Urban [5]	27%	79%
Inside an urbanized area[6]	33%	86%
Outside an urbanized area	67%	14%
% Rural	73%	21%
Rural farm	9%	6%
Rural non-farm	91%	94%

[4] As defined in the 1990 census, the urban population comprises all persons living in urbanized areas or in places of 2500 or more, but excluding those persons living in the rural portion of extended cities. The population *not* classified as urban constitutes the rural population.

[5] It is important that "urban" is not confused with "metropolitan area" in this sense. MSAs or Metropolitan Statistical Areas are individually designated by the Office of Management and Budget and are a core metropolitan area containing a large population nucleus, together with adjacent counties (or towns, in New England) that have a high degree of social and economic integration with that core. MSAs are predominantly urban, but urban areas are not necessarily metropolitan.

[6] The Census Bureau defines 'urbanized area' as one or more places (central place) and the adjacent densely settled surrounding territory (urban fringe) that together have a minimum population of 50,000. The urban fringe usually consists of contiguous territory having a density of at least 1,000 persons per square mile.





Primary among the population characteristics expressed by the phrase 'rural differential', is the density of the population. In rural telco service areas, the number of persons per square mile is 12.87 as compared to 69.31 for the US as a whole, and 104.91 for non-rural telco service areas. Importantly, however, the differences in population density seen among areas served by rural telcos is even more telling. For instance, the mean population density of areas served by rural telcos ranges from .58 persons per square mile in Alaska and 1.25 in Wyoming, to 280 and 247 respectively in Connecticut and New Jersey. To the extent that cost of service is predicated on the density of the population served, the differences shown on a statewide basis--let alone on a study area or wire center basis--are significant indeed.

Table 2: Population Density of Rural and Non-Rural Telephone Service Areas

	Rural Service Areas	Non-Rural Service Areas	U.S.
Persons Per Square Mile	12.87	104.91	69.31

A generally accepted, but unquantified, relationship is that between the population density of areas served by rural and non-rural telcos across states. With the ability to break down population per square mile by telco service area, it can further be seen that significant variations exist among states.

Table 3: Population Density of Rural and Non-Rural Telco Service Areas by State[7]

[7] It should be noted that the percentages of population located in rural and non-rural Telco Service Areas (TSAs) will not in all cases total 100%. All geographic area could not be attributed by TSA. .1% of the US population was located

	% of Total State Pop located in a rural TSA	% of Total Pop located in a non-rural TSA
Alaska	44%	52%
North Dakota	31%	69%
Montana	28%	72%
South Dakota	27%	72%
Iowa	23%	74%
Minnesota	23%	74%
Arkansas	22%	78%
South Carolina	21%	79%
Wisconsin	20%	80%
Pennsylvania	17%	82%
Kentucky	15%	84%
Maine	15%	83%
West Virginia	14%	85%
Tennessee	13%	87%
Oklahoma	12%	87%
Nebraska	12%	88%
Kansas	12%	88%
Vermont	11%	87%
North Carolina	10%	89%
Idaho	10%	89%
Georgia	9%	89%
Alabama	8%	92%
Mississippi	8%	90%
Louisiana	8%	92%
Missouri	7%	92%
New Hampshire	6%	92%
Ohio	6%	94%
Oregon	6%	92%
Arizona	6%	94%
Texas	6%	94%
Washington	5%	94%
Indiana	5%	95%
Illinois	5%	95%
Colorado	5%	93%
Utah	5%	95%
Wyoming	4%	96%
New Mexico	4%	96%
New York	4%	96%
Michigan	3%	96%
New Jersey	3%	97%
Virginia	3%	97%
Nevada	2%	98%
Florida	2%	97%
Connecticut	1%	99%
California	1%	98%
Maryland	<1%	99%
Massachusetts	<1%	99%
Delaware	0%	100%
DC	0%	100%
Hawaii	0%	100%
Rhode Island	0%	100%
US	7%	92%

In twenty states, 10% or more of the population is served by rural telcos. In only three states, i.e.., Rhode Island, Hawaii and Delaware, and the District of Columbia is none of the population served by rural telcos. The regional breakdown among the twenty states, shows that three are located in the West, seven in the Midwest, three in the Northeast and seven in the South, indicating a greater

outside the land boundaries of the U.S, the so-called "island population", and .5% of the population were originally located within wire centers for whom the rural or non-rural status of the telco provider could not be immediately determined. Corrections will which allow for inclusion of the .5% are currently being made.

prevalence of rural telco coverage in the West and Midwest, but certainly not to the exclusion of the Northeast and South.

The density range among states shows that in fourteen states the population of non-rural telco service areas exceeds that of rural TSAs by more than 100 people per square mile. Included among those states are Wisconsin, Indiana, Virginia, Ohio, Michigan, California, Illinois, Florida, Pennsylvania, Maryland, New York, Connecticut, Massachusetts, and New Jersey. These states all share the distinction of being densely populated, and the areas served by rural telcos within those states likewise have a higher population density than other rural TSAs. It is of importance to note, however, that in some very populous states, e.g., California, the density of the rural TSAs is exceedingly low when compared to the density of non-rural TSAs.

Table 4: Comparison of Population and Densities by State

State	1990 State Pop (in millions)	Total Pop Density (Persons per sq. mi.)	Rural Pop Density (Persons per sq. mi.)	Rural TSA Pop Density	Non-Rural TSA Pop Density	Diff
New Jersey	7.7	1042	134	247	1153	-906
Massachusetts	6.0	768	117	154	774	-620
Connecticut	3.3	678	170	280	685	-405
New York	18.0	381	56	88	432	-344
Maryland	4.8	489	83	208	489	-281
Pennsylvania	11.9	265	86	109	381	-272
Florida	12.9	240	33	29	271	-242
Illinois	11.4	206	32	43	253	-210
California	29.8	191	14	8	211	-203
Michigan	9.3	164	29	22	211	-189
Ohio	10.8	272	68	124	287	-163
Virginia	6.1	156	47	39	171	-132
Indiana	5.5	155	56	54	173	-119
Wisconsin	4.9	90	26	34	153	-119
South Carolina	3.5	116	52	60	153	-93
Georgia	6.5	112	42	42	132	-90
Tennessee	4.9	118	48	57	142	-85
Washington	4.9	73	17	14	97	-83
Louisiana	4.2	97	27	35	113	-78
Texas	17.0	65	13	12	90	-78
Missouri	5.1	74	24	20	96	-76
New Hampshire	1.1	124	62	61	135	-74
Minnesota	4.4	55	16	24	91	-67
Kentucky	3.7	93	45	47	113	-66
North Carolina	6.6	136	64	82	148	-66
Oklahoma	3.1	46	15	13	73	-60
Iowa	2.8	50	20	22	80	-58

State	1990 State Pop (in millions)	Total Pop Density (Persons per sq. mi.)	Rural Pop Density (Persons per sq. mi.)	Rural TSA Pop Density	Non-Rural TSA Pop Density	Diff
Alabama	4.0	80	32	35	91	-56
Colorado	3.3	32	6	3	56	-53
Arizona	3.7	32	4	5	51	-46
Maine	1.2	40	20	13	54	-41
Kansas	2.5	30	9	8	47	-39
West Virginia	1.8	74	48	45	83	-38
Arkansas	2.4	45	21	25	59	-34
Vermont	0.6	61	40	36	70	-34
Mississippi	2.6	55	29	28	59	-31
Nebraska	1.6	21	7	5	36	-31
Utah	1.7	21	3	3	30	-27
Oregon	2.8	30	9	10	35	-25
New Mexico	1.5	12	3	2	15	-13
South Dakota	0.7	9	5	4	17	-13
Idaho	1.0	12	5	4	16	-12
Nevada	1.2	11	1	2	12	-10
North Dakota	0.6	9	4	5	15	-10
Montana	0.8	5	3	3	8	-5
Wyoming	0.5	5	2	1	5	-4
Alaska	0.6	1	<1	0.58	1.8	-1.2
Delaware	0.7	341	79	-	341	-
DC	0.6	9951	-	-	9880	-
Hawaii	1.1	173	12	-	172	-
Rhode Island	1.0	960	118	-	960	-
US	248.7	70	17	13	105	-92

While not a revelation, it is clear from the above table, that in every state (in which there are rural telcos operating) the density of the rural telco service areas falls significantly below that of non-rural TSAs.

Also of interest is the correlation found between the density of non-rural TSAs across all states and the total population density. Through the computation of a product-moment correlation coefficient, it was determined that a near-perfect correlation existed between the two densities ($r = .9998$) across all states, thereby signifying the overwhelming extent to which the "average" exchange served by non-rural telcos in a state corresponds to the average density for the state as a whole. When doing a similar correlation analysis between the density of rural TSAs and the density of each state's census-defined rural population as a whole, a very high correlation was also found ($r = .9545$), further indicating that the "average" population served by rural telcos within each state is almost exclusively rural, isolated, and/or remote. This state-level comparison between the rural and non-rural TSAs further corroborates the need for a universal service mechanism that offsets the much higher cost of providing telecommunications service to low-density populations.

Table 5: Percentage Difference Between Population Densities of Rural and Non-Rural Telco Service Areas (TSAs) by State

State	Region	Population Density of Non-Rural TSAs	Population Density of Rural TSAs	% Difference
North Carolina	S	148	82	45%
West Virginia	S	83	45	46%
Vermont	NE	70	36	49%
Mississippi	S	59	28	53%
New Hampshire	NE	135	61	55%
Ohio	MW	287	124	57%
Maryland	S	489	208	57%
Arkansas	S	59	25	58%
Kentucky	S	113	47	58%
Connecticut	NE	685	280	59%
Tennessee	S	142	57	60%
South Carolina	S	153	60	61%
Alabama	S	91	35	62%
Montana	W	8	3	63%
North Dakota	MW	15	5	67%
Alaska	W	1.8	0.58	68%
Georgia	S	132	42	68%
Indiana	MW	173	54	69%
Louisiana	S	113	35	69%
Pennsylvania	NE	381	109	71%
Oregon	W	35	10	71%
Iowa	MW	80	22	73%
Minnesota	MW	91	24	74%
Idaho	W	16	4	75%
Maine	NE	54	13	76%
South Dakota	MW	17	4	76%
Virginia	S	171	39	77%
Wisconsin	MW	153	34	78%
New Jersey	NE	1153	247	79%
Missouri	MW	96	20	79%
New York	NE	432	88	80%
Wyoming	W	5	1	80%
Massachusetts	NE	774	154	80%
Oklahoma	S	73	13	82%
Kansas	MW	47	8	83%

Illinois	MW	253	43	83%
Nevada	W	12	2	83%
Washington	W	97	14	86%
Nebraska	MW	36	5	86%
New Mexico	W	15	2	87%
Texas	S	90	12	87%
Florida	S	271	29	89%
Michigan	MW	211	22	90%
Utah	W	30	3	90%
Arizona	W	51	5	90%
Colorado	W	56	3	95%
California	W	211	8	96%
US		105	13	88%

Table 5 above shows the percentage difference in population density among states between rural and non-rural TSAs. The state showing the highest variation between the population density of rural vs. non-rural TSAs is California in which there is a 96% difference. At the other end of the range is North Carolina, West Virginia and Vermont for whom the percentage difference between rural and non-rural TSA densities falls between 45-49%. In all but those three states, the mean population density of rural TSAs exceeds that of non-rural TSAs by more than 50%.

Race

The racial composition of areas served by rural telcos differs markedly from that of non-rural service areas, as can be seen in Table 6. The minority population of non-rural telco service areas accounts for approximately 20% of the population, roughly twice that in rural telco areas. Notable exceptions exist, however. Rural telco service areas in Arizona and South Carolina lead the country in percent of minority populations with 46.6% and 33.2% respectively. Across non-rural telco service areas, the states in which the highest racial minority population resides are the District of Columbia (70.4%) and Hawaii (66.6%).

Among rural telco areas, Asian and Pacific Islanders account for 2.2% of Alaskan citizens and 2.1% of California citizens, while 61.9% of Hawaiians and 9.6% of Californians represent the highest level of Asian and Pacific Islanders within non-rural telco service areas.

American Indian, Eskimo and Aleut populations are largest among rural telco areas in Arizona (43.4%) and in Alaska (19.8%), but among non-rural telcos areas, the largest native American populations occur in Alaska (11.7%) and New Mexico (8.5%). It should also be noted that while the Native American population of rural TSAs in Arizona comprise nearly half of their population base, the Native American population of non-rural TSAs--while slightly larger in number--makes up less than 3% of the population base of non-rural telcos.

Unlike other minority groups, Hispanic populations do not differ significantly between rural and non-rural telco service areas. Persons of Hispanic origin account for 36.8% of those served by rural telcos in New Mexico and 24.2% of those served by rural telcos in Texas. Similarly, 38.1% of the persons served by New Mexico non-rural telcos are Hispanic, with California ranking as the next highest state (25.5%) in terms of Hispanic customers of non-rural telcos. Differences exist, however, in the ethic diversity among specific service areas.

Table 6: Racial Composition of Rural and Non-Rural Telephone Service Areas

	Rural Service Areas	Non-Rural Service Areas	
White	90.5%	79.6%	White
Black	5.6%	12.5%	Black
Asian & Pacific Islander	.5%	3.1%	Asian & Pacific Islander
American Indian, Eskimo and Aleut	2.1%	0.7%	American Indian, Eskimo and Aleut
Outlier States – High Minority			Outlier States – High Minority
Arizona	46.6%	70.4%	District of Columbia
South Carolina	33.2%	66.6%	Hawaii
Outlier States – Low Minority			Outlier States – Low Minority
Maryland / New Hampshire	.8%	1.5%	Vermont
Iowa	.9%	1.7%	Maine
Outlier States – High Black Population			Outlier States – High Black Pop.
South Carolina	32.2%	65.9%	District of Columbia
Mississippi	30.4%	36.5%	Mississippi
Outlier States – High Asian Population			Outlier States – High Asian Pop.
Alaska	2.2%	61.9%	Hawaii
California	2.1%	9.6%	California
Outlier States – High Native Am. Pop.			Outlier States – High Native Am.
Arizona	43.4%	11.7%	Alaska
Alaska	19.8%	8.5%	New Mexico
Outlier States – High Hispanic Pop.			Outlier States – High Hispanic
New Mexico	36.8%	38.1%	New Mexico
Texas	24.2%	25.5%	California

Focusing on the native American population across states, it can be seen in Table 7 below that in all but one of the fifteen states in which the Native American population comprises 1.5% or more of the population residing in rural TSAs, the Native American population represents a significantly larger percentage of the population base of rural telcos. In Arizona alone more than 40% of the population base of rural telcos is made up of Native Americans, as compared to about 3% of the population base for non-rural telcos. This is of primary importance in discussions of universal service, given the severe problems with Native American telephone penetration and the sparseness of the rural telco service areas in which the Native American population resides.

Table 7: Percentage of Native American Population by Telco Service Area for Selected States

State	Region	Native Am Pop as % of Total Pop Served by Non-Rural Telcos	Native Am Pop as % of Total Pop Served by Rural Telcos
Arizona	W	3.1%	43.4%
Alaska	W	11.7%	19.8%
New Mexico	W	8.5%	17.7%
South Dakota	MW	4.2%	15.7%
Oklahoma	S	7.3%	12.9%
Montana	W	5.3%	7.7%
California	W	0.8%	4.0%
North Dakota	MW	4.0%	3.9%
Nevada	W	1.7%	3.3%
Washington	W	1.6%	2.9%
Utah	W	1.4%	2.7%
Idaho	W	1.3%	2.5%
Nebraska	MW	0.6%	2.1%
Michigan	MW	0.6%	1.8%
Colorado	W	0.8%	1.5%
U.S.		0.7%	2.1%

Type of Household

Areas served by rural telcos are more likely to have a higher percentage of family households, e.g., 76% of all households are comprised of families in rural telco areas as compared to 70.4% in non-rural telco areas. Likewise, married couples comprise the households in 85.1% of rural telco areas, but in only 79% of non-rural telco areas. The portion of households comprised of married couples range from a high of 89% in Nebraska and Utah to a low of 78% in Arizona and South Carolina. Similar state comparisons across non-rural telco areas show a high of 85% in Idaho, North Dakota, and South Dakota and a low of 52% in the District of Columbia.

Female-headed households account for approximately 11% of households in rural telco areas and more than 16% in non-rural telco areas. The range of percentages across rural telco areas by state was not as wide--6.9% in Wyoming to 17.4% in South Carolina--as in those areas served by non-rural telcos--10.8% in Idaho to 39.2% in the District of Columbia.

Non-family households were somewhat less likely to be found in those areas served by rural telcos, i.e., 24.0%, as compared to 29.6% in non-rural service areas.

Table 8: Household Composition in Rural and Non-Rural Telephone Service Areas		
	Rural Service Areas	Non-Rural Service Areas
Family Households	76.0%	70.4%
Households with Married Couples	85.1%	79.0%
With Own Children	40.8%	37.0%
Female-Headed Households	11.1%	16.3%
With Own Children	6.3%	9.2%
Non-Family Households	24.0%	29.6%

Income

Generally, the population in areas served by rural telcos earns significantly less when evaluating any measure of income. Interestingly, however, while the percentage of persons served by rural telcos who fall below the poverty level is somewhat higher than in areas served by non-rural telcos, the percentage of persons falling at or below 50% of the poverty level is similar.

Table 9: Household Income in Rural and Non-Rural Telephone Service Areas		
	Rural Service Areas	Non-Rural Service Areas
Less than $10,000	18.6%	15.2%
$10,000 - $24,999	30.9%	26.0%
$25,000 - $49,999	34.2%	33.7%
$50,000 - $99,999	14.3%	20.6%
$100,000 and over	2.9%	4.6%
Median Household Income	$25,282	$30,418
Average Household Income	$31,211	$38,983

Table 10: Other Income Measures in Rural and Non-Rural Telephone Service Areas		
	Rural Service Areas	Non-Rural Service Areas
Median Family Income	$29,643	$35,728
Average Family Income	$35,247	$44,475
Per Capita Income	$11,379	$14,611

% Persons Below Poverty Level	14.6%	13.0%
% Persons Below 50% of Poverty	5.7%	5.8%

On nearly every measure of income the range across states among persons served by rural telcos exceeds that for those served by non-rural telcos. For example, in 1990 the median family income among persons served by rural telcos ranged from a state low of $21,112 in Kentucky to a state high of $59,133 in New Jersey. Among persons served by non-rural telcos, median family income ranged from a state low of $24,296 in Mississippi to a high of $49,082 in Connecticut. Keeping in mind that the median family income is more than $6000 less in rural than in non-rural service areas, this corroborates the importance of the differences *among* rural service areas and the potential impact of increases in basic phone rates.

Educational Levels

The educational level of persons residing in rural telco areas is lower than in non-rural areas, as indicated by the 14% and 10%, respectively, of persons with less than a 9th grade education as shown in Table 11. Terminal high school diplomas (or their equivalent) were more prevalent in rural telco service areas where 36% of the population had achieved a high school diploma as compared to 29.5% in non-rural service areas. Non-rural telco areas exceeded rural telco areas, however, in the percentage of persons attending some college (19.0% vs. 15.7%), holding Associate and Bachelor's Degrees (19.6% vs. 14.3%), and gaining graduate or professional degrees (7.4% vs. 4.2%)

Table 11: Educational Levels in Rural and Non-Rural Telephone Service Areas

	Rural Service Areas	Non-Rural Service Areas
% With Less than 9th Grade Education	13.9%	10.1%
% With 9th-12th Grade Education (No Diploma)	15.6%	14.3%
% High School Graduate (or equivalent)	36.3%	29.5%
% Some College (No Degree)	15.7%	19.0%
% Associate / Bachelor's Degree	14.3%	19.6%
% Graduate or Professional Degree	4.2%	7.4%

Occupations

As the 1990 Census data included in this analysis do not reflect the more recent escalation of high-tech jobs, the occupational trends outlined below significantly underemphasize the growing need for access to advanced telecommunications services in areas served by rural telcos. However, as can be seen in Table 12, persons residing in rural service areas are much more likely to be employed in farming, forestry, and fishing (7.1%) than are persons residing in non-rural telco service areas (2.1%). But the single largest category of occupations in rural telco service areas is "other"

(34.3%), pointing to the occupational diversity of persons residing in those areas. Regardless of the occupational categories to which rural TSA residents belong, trends point to the increasing occupational requirements for access to advanced telecommunications services.

Non-rural service areas show a significantly larger percentage of residents employed in managerial and professional occupations (27.0% vs. 19.6%) and in technical, sales, and administrative support (32.1% vs. 25.9%). Only in service occupations are the two areas roughly equivalent in terms of the percentage of persons employed (13.2% vs. 13.1%)

Table 12: Occupations in Rural and Non-Rural Telephone Service Areas

% of Persons Age 16 and Older Who are Employed in:	Rural Service Areas	Non-Rural Service Areas
Managerial and Professional Occupations	19.6%	27.0%
Technical, Sales, and Administrative Support Occupations	25.9%	32.1%
Service Occupations	13.1%	13.2%
Farm, Forestry, and Fishing Occupations	7.1%	2.1%
All Other Occupations	34.3%	25.6%

Housing Units

In 1990 more than 15% of all housing units were vacant in areas served by rural telcos as compared to under 10% in areas served by non-rural telcos. Persons residing in rural telco areas are more likely to own the home in which they reside and less likely to rent their home. As with other demographic characteristics, the differential expressed across states is dramatic. Among rural telco service areas, the percentage of owner-occupied units ranges from a low of 44.9% in Alaska to a high of 78.6% in Maryland. Among non-rural telco areas, it ranges from a low of 34.9% in Washington, D.C. to a high of 65.1% in West Virginia.

The prevalence of mobile homes in rural telco areas (17.6% of all homes) is more than double that in non-rural telco areas (7.2%).

Rental properties, whether they be homes, apartment complexes, or condominiums are much less likely to be found in areas served by rural phone companies (20.1% vs. 33.1% in non-rural TSAs).

Table 13: Type of Housing Units in Rural and Non-Rural Telephone Service Areas		
	Rural Service Areas	Non-Rural Service Areas
% Occupied Units	84.4%	90.4%
% Owner-Occupied Units	64.3%	57.3%
% Rented Units	20.1%	33.1%
% Vacant Units	15.6%	9.6%
% Single Family Units	74.5%	63.5%
% Buildings with 5+ Units	4.7%	18.8%
% Condominiums	1.4%	5.5%
% Mobile Homes	17.6%	7.2%

Residents of rural TSAs are more likely to own their homes than are their counterparts in non-rural TSAs, but little difference exists across TSAs in the age distribution of householders.

Table 14: Householder Status and Age Distribution of Householders In Rural and Non-Rural Telephone Service Areas		
	Rural Service Areas	Non-Rural Service Areas
% Owner-Occupied Households	76.1%	63.4%
Age of Owner – Householders		
% 15-64	73.0%	74.0%
% 65-74	15.7%	15.7%
% 75+	11.3%	10.3%
% Renters	23.9%	36.6%
Age of Renters		
% 15-64	82.8%	84.6%
% 65-74	7.9%	7.6%
% 75+	9.3%	7.8%

Housing Values

The average home value in areas served by rural telcos was $67,456 (based on 1990 census data) as compared to $113,889 in non-rural service areas. Median home values showed less variance, but still the homes served by rural telcos were on average more than $26,000 below the value of those located in non-rural telco service areas.

Table 15: Housing Values in Rural and Non-Rural Telephone Service Areas		
	Rural Service Areas	Non-Rural Service Areas
Average Value of Owner-Occupied Homes	$67,456	$113,889
Average Rent Paid	$359	$494
Median Home Value	$54,878	$80,883
Median Rent Paid	$333	$452

Telephone Penetration

As part of the 1990 Census, long-form respondents were asked whether there was a telephone present in the home. At that time, 94.9% residing in non-rural telco areas and 92.5 % residing in rural telco areas indicated that a telephone was located in the housing unit. While telephone penetration rates have increased slightly since 1990, this data is included here because of the detail available with respect to owner/renter status and age of householder.

More than three-fourths (78.6%) of housing units with telephones in areas served by rural telcos are owner-occupied homes as compared to rental units (21.4%). This ratio of owner-occupied to renter population is much different in areas served by non-rural telcos, where only 65.4% of the housing units with telephones are owner-occupied homes while 35.6% are rental units.

Perhaps of more interest is the difference between homeowners and renters *without* telephones across TSAs. Those persons without phones who reside in rural TSAs are much more likely to own their own home (46.0%) than to rent (54.0%). Conversely, non-rural TSA residents without phones are much more likely to rent their home (75.2%) than to own it (24.8%).

Table 16: Extent of Telephone Penetration in Rural and Non-Rural Telephone Service Areas		
	Rural Service Areas	Non-Rural Service Areas

% of Housing Units with Telephone	92.5%	94.9%
% With Telephone who Own Their Home	78.6%	65.4%
% With Telephone who Rent	21.4%	34.6%
Age of Householder with Telephone		
15-59	66.7%	70.2%
60-64	7.6%	7.1%
65-74	14.4%	13.1%
75+	11.3%	9.6%
% of Housing Units with no Telephone	7.5%	5.1%
% Without Telephone who Own Their Home	46.0%	24.8%
% Without Telephone who Rent	54.0%	75.2%
Age of Householder with no Telephone		
15-59	81.9%	85.2%
60-64	4.9%	4.2%
65-74	7.6%	6.4%
75+	5.6%	4.2%

Viewing the data in a slightly different way, we can infer from Table 17 that homeowners are more likely to have telephones regardless of whether they are located in a rural or non-rural TSA. Similarly, older householders are more likely to have telephones regardless of the telephone service area in which they reside. However, householders in rural TSAs, regardless of age or home ownership, are less likely to have phones, than are their counterparts in non-rural TSAs.

Table 17: Age and Tenure of Householders with Phones in Rural and Non-Rural Telephone Service Areas				
	Rural Service Areas		Non-Rural Service Areas	
Age of Householder	With Phones	W/O Phones	With Phones	W/O Phones
% 15-64	91.4%	8.6%	94.2%	5.8%
% 65-74	95.9%	4.1%	97.4%	2.6%
% 75+	96.1%	3.9%	97.7%	2.3%
Tenure of Householder	With Phones	W/O Phones	With Phones	W/O Phones
Owner-Occupied	95.5%	4.5%	98.0%	2.0%
Renter-Occupied	83.1%	16.9%	89.6%	10.4%

In 1990, there were 4,313,692 households (5.1%) in areas served by non-rural telcos that had no telephone. In areas served by rural telcos, the number of households that had no phone was 469,537 (7.5%). Of those 4,783,229 householders having no telephone, 90% were located in areas served by non-rural telcos, while 10% were located in rural telco service areas.

Table 18: % of Households by TSA Location and Telephone Status		
	Rural Service Areas	Non-Rural Service Areas
% of households located in TSA	6.9%	93.1%
% of households without phones	7.5%	5.1%

Summary of Other Demographic Characteristics within Rural Telephone Service Areas

The following table highlights other demographic characteristics of the population residing within rural TSAs across the U.S:

Table 19: Demographic Summary of Population Residing in Rural TSAs

Age	
Under 17	28%
18-49	45%
50-64	14%
65-79	10%
Over 80	3%
Sex	
Male	49%
Female	51%
Language Spoken in Household	
English	91%
Spanish	3%
Asian / Pacific Island	<1%
Other	5%
Place of Birth	
Born in State of Residence	74%
Born in Other U.S. State	23%
Northeast	4%
Midwest	8%
South	8%
West	<4%
Born Outside U.S.	<2%
Foreign Born	2%
Place of Work (Workers over Age 16)	
In county of residence	72%
Outside county of residence / in-state	24%
Outside county of resident / out of state	4%
Means of Transportation to Work	
Car/truck/van – drove alone	74%
Carpooled	15%
Walked	4%
Worked at home	5%
Other	<2%
Class of Worker (Employed Persons Over 16)	
Private Profit Workers	68%
Private Not-for-Profit Workers	6%
Local/State/Fed Government Workers	15%
Self-Employed Workers	11%
Unpaid Family Workers	<1%

Conclusion

Endemic to the discussion of current telecommunications issues and the debates surrounding universal service is the extent to which rural telephone companies and the populations which they serve are markedly different from other LECs (Local Exchange Carriers) or CLECs (Competitive Local Exchange Carriers). This debate will undoubtedly continue as the support mechanisms for universal service are challenged and/or altered to better accommodate a competitive telecommunications industry. Of equal relevance to this discussion, are the verifiable differences, as laid out in this paper, which exist across rural company service areas.

The foregoing analysis was conducted largely as a precursor to the forthcoming 2000 census, so that the methodology could be perfected well in advance of collection and release of the millennial census data. It is fully acknowledged that much of the data that resulted from this analysis may not be current. It is anticipated, however, that the differences which exist between a breakdown of 1990 and 2000 datasets across rural and non-rural TSAs will, in most instances, be largely one of degree.

It has been shown that the population served by rural telephone companies is markedly different than that of non-rural companies with respect to many demographic characteristics. Of even more impressive proportions are the differences which exist across rural telco service areas. For individual rural companies, the ability to ascertain differences in population demographics down to the level of the wire center can be an important tool in understanding the need and explaining the case for the differential handling of universal service mechanisms across rural and non-rural companies. Finally, not to be overlooked is the ability which access to such data can mean for rural telcos in terms of financial planning, market growth analysis, and anticipated demand for advanced services.

Part II: Methodology
Allocating Census Summary Data to Telephone Exchanges

The Problem

The need arose for detailed population and housing data from the 1990 decennial census summarizing the geographic areas that correspond to telephone exchange service areas. Such service areas are linked to a physical geographic entity called a Wire Center. There is a many-to-one link between telephone exchange areas and the Wire Centers. The goal, therefore, was to aggregate the data to the Wire Center level. Once data was derived for the Wire Center ("WC") it would be possible to start with any telephone exchange area (consisting of 3-digit area code and a 3-digit dialing prefix) and relate it to a Wire Center -- and thus to the census tables summarizing that WC. Since the Census Bureau does not publish data for WC's, it was necessary to create it by aggregating the summaries for the smallest geographic entities for which the data were available. This gave rise to questions as to how such small census geography units could be related to the WC's, and whether these units were small enough so that it was feasible to aggregate (or disaggregate) them to the Wire Center level.

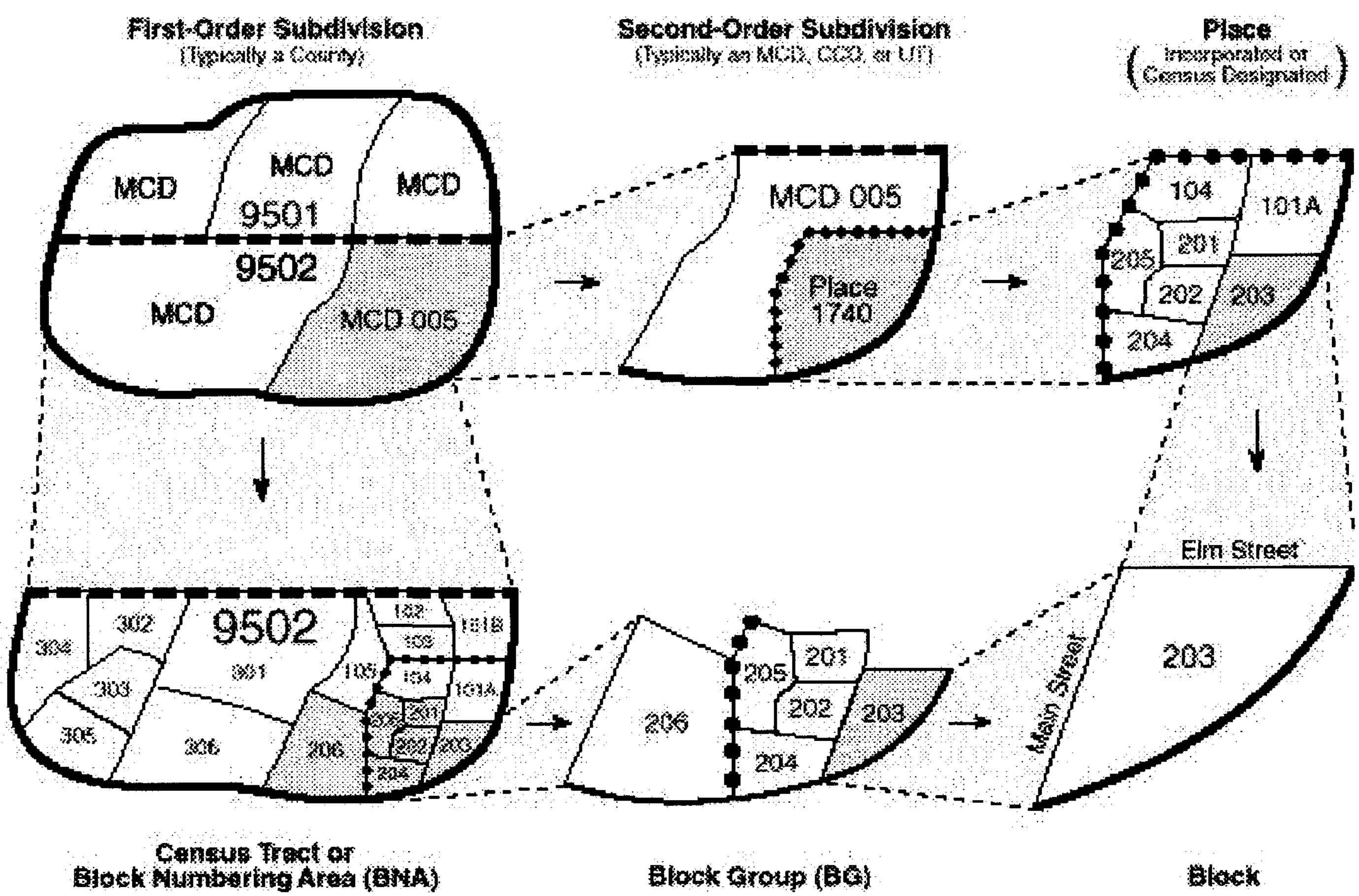
Once the data was merged into Wire Center units, we were then interested in being able to determine how those Wire Centers related to rural telephone service areas. A number of data sources allowed us to estimate these relationships. The common key was the "npa-nxx" codes -- or what we called the "telephone exchange" (3-digit dialing prefix within area code within state.) From our data vendor (On Target Mapping) we had a file that provided a set of telephone exchanges that corresponded to each wire center. From NECA we were able to obtain a set of files that defined the specific telephone companies, which were designated as having rural service areas, as well as files that could be used to link telephone exchanges with the companies. Combining all these sources would permit us to determine the approximate proportion of each Wire Center that was associated with one of the rural service providers. We would use this correspondence to do another data allocation to apportion Wire Center data to what we called the "Rural Service Area" portions of each state and the U.S.

The Tools

Census Geography Basics

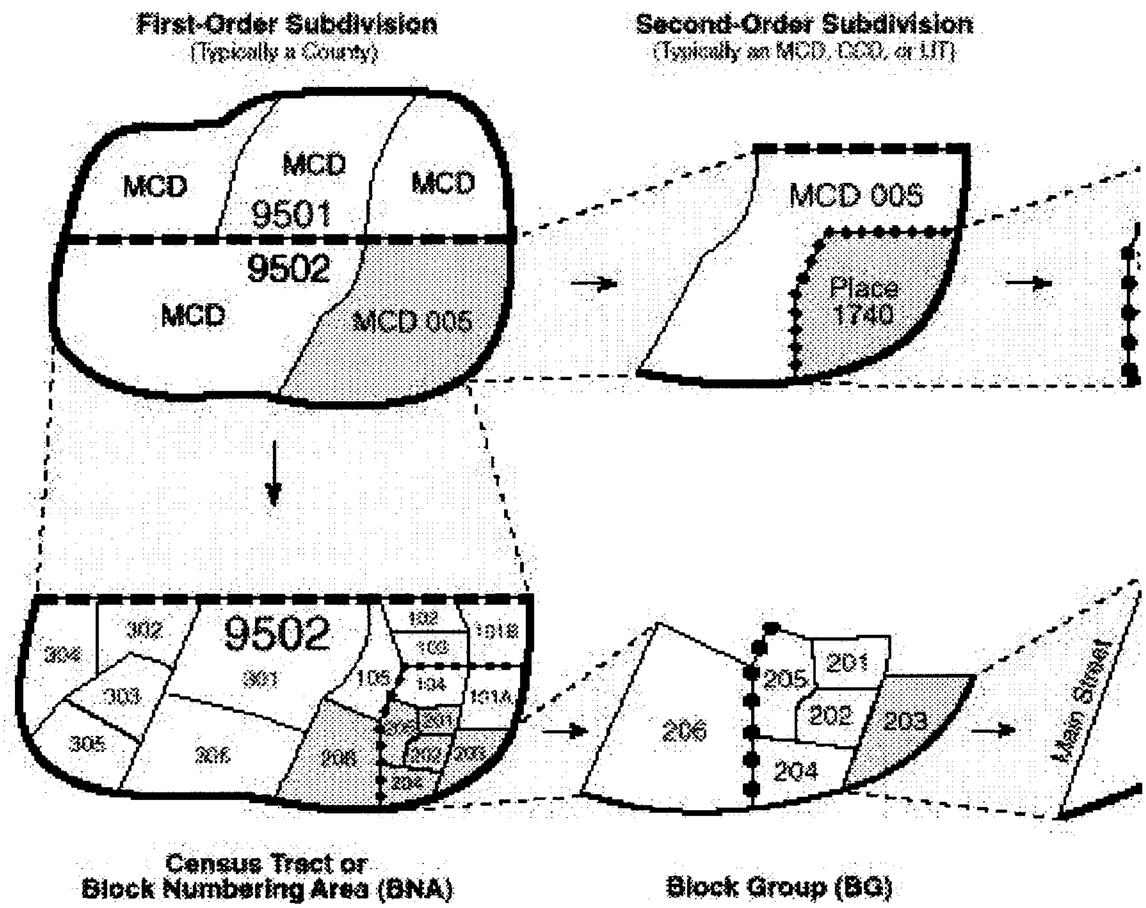
The Census Bureau uses a very complex geographic scheme for tabulating the results of the decennial census. It is so complex that they have published a several hundred-page book called the ***Geographic Areas Reference Manual.*** This book is now accessible via the web at http://www.census.gov/geo/www/garm.html. Included in that document (a collection of Adobe Acrobat PDF files) is a graphical depiction of the "Small Area Geography" used in the 1990 Census (see below). The geographic scheme it describes is a complex hierarchy that starts at the county level. Within the county are units called *MCD*'s (Minor Civil Divisions, also sometimes referred to as "County subdivisions" - in many states they are also known as "towns" or "townships"). MCD's are then subdivided by *Places*, a generic term used by the Census Bureau to refer to local governmental units. For the purposes of this scheme the "unincorporated remainder" of a MCD is also considered a (pseudo) place. Continuing down after the place level we have census tracts (or Block Numbering Areas in most rural areas -- same thing for most purposes), block groups and, at the very bottom of the hierarchy, **census blocks**. An important fact to know about these census blocks is that for the purposes of defining geographic entities to be used in tabulating the census, *a block is an indivisible or "atomic" unit*. All other geographic units used to tabulate the census can be described as a union of census blocks.

Figure 2-3. **Small Area Geography in the 1990 Census**



And, most important, a block will never split any of the other units in the census scheme. So if you have data on the number of persons by age, race and sex tabulated by census blocks, you know you can aggregate that data to get data by place, MCD, tract, etc. So census blocks are extremely important as geographic building blocks. There are just over 7 million of these blocks nationwide. (These are the blocks used in the *1990* census - there is a different set used each decade.)

Figure 2-3. **Small Area Geography in the 1990 Census**



Block Headers File

The Census Bureau has provided census data users with a special collection of geographic reference files (known as the "*Block Headers*" files, one file per state) that contain information about each of these seven million blocks. Each record contains data for one block and contains all the other related geographic codes – state, county, MCD, place and tract codes, for example. It also contains the 1990 complete count population of the block, the urban/rural classification of the block (since the Bureau tabulates data for the urban / rural portions of geographic areas, each block must be entirely one or the other.) Finally, each record in the Block Headers file contains the latitude, longitude coordinates of a point located within the boundaries of the block (usually the spatial centroid, but sometimes if that centroid falls outside the block boundary then the point is moved to a nearby internal point.) All of the information in these files is derived from the Bureau's TIGER geographic reference database. There are no typos and no unverified coordinates. It is an official and complete description of what blocks exist and how they relate to all the other census geographic entities.

A special version of the Block Headers file was created with the addition of 5-digit ZIP codes to the files. The details of how and why this was done are not relevant to this application so we won't deal with them here, except to say that they resulted in something called the *ZIP Headers* files, which are the actual source files used in generating the MABLE database, discussed below.

MABLE/Geocorr Web Application

Using these *ZIP Headers* files the author, working under a contract with the Consortium for International Earth Sciences Network (CIESIN), created a special database and corresponding web application called ***MABLE/Geocorr***. MABLE – an acronym for Master Area Block Level Equivalency – was the database, and Geocorr was the engine (program) that ran as a web application and allowed users to manipulate the data in MABLE to create geographic correspondence files. It allowed users to see the degree of correspondence (geographic intersection) between two sets of geographic codes (the "source" and "target" layers.) For example, the user could specify that they wanted to look at county and census tract as the source codes, with place (city) and ZIP codes as the target layer. They could specify a geographic universe for the reports corresponding to one or more states, counties, metro areas or cities. Shown below is a sample of a portion of the output of such a request for a single county (Audrain, Missouri.)

COUNTY	Census Tract/BNA	PLACE CODE (FIPS)	ZIP CODE	Place Name	1990 POP	Tract to ZIP alloc factor
29007	9501	75688	63382	Vandalia	2683	0.914
29007	9501	99999	63382		252	0.086
29007	9502	4834	65232	Benton City	139	0.027
29007	9502	23662	63345	Farber	418	0.081
29007	9502	39602	63352	Laddonia	575	0.112
29007	9502	39602	65280	Laddonia	6	0.001
29007	9502	46460	65264	Martinsburg	337	0.066
29007	9502	63560	65280	Rush Hill	121	0.024
....	....	....	....	....	....	....
29007	9503	99999	65285		1187	0.433
29007	9504	47648	65265	Mexico	2864	0.951
29007	9504	75706	65265	Vandiver	75	0.025
29007	9504	99999	65265		72	0.024
29007	9505	47648	65265	Mexico	3057	0.815
29007	9505	99999	65265		693	0.185
29007	9506	47648	65265	Mexico	2048	0.827
29007	9506	99999	65265		429	0.173
29007	9507	47648	65265	Mexico	3321	0.937
29007	9507	99999	65265		225	0.063

Look at the highlighted rows (near the end of the above table) for Tract/BNA 9504. By selecting all the blocks in that BNA and keeping the place and ZIP codes that went with them, the *geocorr* program is able to tell us how many people (in the specified county) were living in all the BNA-place-ZIP combinations in 1990. 75 people, or about 2.5% of the total population of 9504, also lived in the town of Vandiver and the 65265 ZIP code. 95.1% of the population lived within the city limits of Mexico and the 65265 ZIP code. Another 72 people lived in pseudo-place 99999 (unincorporated remainder of the county) within ZIP 65265.

<u>Summary Level 090</u>

What we just did for relating census tract/BNA's to place-ZIPs for one county we can do for *any* two sets of geographic code combinations stored in the MABLE database. For example (and for relevance to our Wire Center data problem), we could select for our set of "source" geocodes: county, tract/BNA, MCD, place, urbanized area, urban/rural part, congressional district (102^{nd}), and block group. For the target area we could select any other code, e.g. hydrological units (watersheds). *Why* would we want to do such a thing? Because this specific seemingly unlikely combination of source geographic codes represents a special geographic "summary level" used in tabulating the 1990 census. Every census questionnaire is "geocoded" by the Census Bureau to identify all these various geographic units. The data that gets released to the public is mostly in the form of "Summary Tape Files". These STF's consist of pre-defined tables that summarize the people and housing units within geographic entities formed by taking various *combinations* of geographic codes. *The smallest geographic unit for which the Bureau publishes tabulations based on the "Long Form" (1-in-6 sample) data is this particular list of codes.* (Data based on the short form, also known as *complete count* data, is available down to the block level.)

Each combination of geographic codes for which the Bureau publishes summary data is assigned a 3-character code. These codes are included on the summary files ("STF"s) to let users select the type of geographic summary they want. The code for our special smallest-unit-for-sample-data combination is "090". So we refer to data summarized for this combination as "090 level" data, or sometimes as "split block group" data. Census block groups are unique within tract/BNA within county, and data for unsplit block groups (summary level code "150") are commonly used and sometimes thought to be the smallest unit for which sample tabulations are available. But the 090 level will provide more detailed data whenever an MCD, place, CD, urbanized area, or urban-rural part splits the block group.

Data Aggregation/Allocation

A very common task for people who work with census (and other) pre-aggregated data is to take data summarized by one set of geographic units and to recast those summaries into other geographic entities for which summary data is needed. To do this requires:

- A geographic equivalency file that defines the relationship between the geographic units for which we *have* summary data (the "*source*" units) and those for which we want to allocate it (the "*target*" units). Sometimes called a *correlation list* (or equivalency file, crosswalk file, etc.).
- A program that will combine ("join") the tabulations for the source geography with the correlation list and retabulate the data to the target geographic units.

When a source geography unit is not completely located within a single target geographic unit (for example, when a split block group is in more than one Wire Center) then the program that does the allocation/aggregation actually does *disaggregation* of the tabular data before reaggregating it to the target units. The correlation list needs to provide an "allocation factor" to indicate the estimated portion of the source area that is to be allocated to each of the corresponding target units. The sum of these allocation factors for each source unit must always be 1 (look back at our example for Audrain County to see how all the allocation factors sum to 1 for each tract/BNA.) All tabular counts are multiplied by these allocation factors and written to an intermediate file of source/target summary levels (e.g. a split block group within Wire Center.) The intermediate file is then sorted by the target geography and aggregated so that the final output has only one record/observation per target geographic unit. Note that this process is making the assumption that we can uniformly allocate the source data to the target areas, which may not always be a very good assumption. What we want to avoid is doing allocations where the source geography is larger than the target geography, requiring a lot of this uniform allocation (disaggregation). For example, attempting to allocate county level data to ZIP codes would not produce a very reliable output file since counties are generally much larger than ZIP codes and so the allocation factors would tend to be small. More important, the assumption that characteristics are homogeneously distributed to ZIP codes within a county is often blatantly false. Going the other way, however–allocating/aggregating ZIP-level data to counties–would generally work just fine.

Aggregating Aggregate Data

Special care needs to be taken when working with pre-aggregated data to allocate/aggregate it to new summary levels. Specifically, one needs to be careful when dealing with items such as means, averages, ratios, percentages and medians. None of these items can be simply summed up as you can with data representing counts. If I am aggregating a table of persons by age and I have to allocate the data to give 80% of it to one target unit and 20% of it to another, it is a simple matter of multiplying each of the cell counts by .8 and .2, respectively and giving those counts to the respective target units. But if I have a variable on the input file that has the average age, it makes no sense to take that average and multiply it by .8 and .2 and then to reaggregate these numbers to get a new average age for the target units. These items need to be weighted by their "universe variables" (a variable on the input dataset that measures the size of the population being averaged), aggregated and then recalculated after aggregation to get a weighted average for the output data. This yields not just a good estimate, but an *exact value* on the output data file. For the average age example, I would weight that variable using a variable containing the total population. If I was allocating an average value of selected owner-occupied housing units, then I would need to weight it using a variable that contained the count of those selected owner-occupied units.

Percentages can also be handled as weighted averages. If I have a variable representing the percent of population that is white, I can weight that item by the total population and the weighted average will yield the exact percentage white on the output aggregated file.

Medians are more difficult to aggregate. If you do not have a distribution table to correspond to the median but you do have a universe variable (i.e. if you have the median age of single mothers but you do not have a table showing the age distribution of single mothers but you do have a variable with the count of single mothers) then you can do an unreliable estimate by taking a weighted average of the median. This does not work well for skewed distributions. By far the preferred method is to aggregate a corresponding distribution table and then use a standard statistical method for estimating the median of an interval-based distribution. To see an example of such a method, written in SAS, see the source of the median macro at:
http://www.oseda.missouir.edu/mscdc/sasmacro/median.sas.

The Agg Macro

Because the logic for doing such aggregating of aggregate data is somewhat tedious and error prone, the Urban Information Center at the University of Missouri St. Louis (where the author once worked for many years) developed a special SAS macro for handling the task. The macro has parameters to allow specifying lists of means and their corresponding weight/universe variables. It also accepts a parameter indicating there is a variable on the input data set that is to be used as an "allocation factor" so that the numbers in this observation get multiplied by this factor prior to being added to the totals. This macro can be viewed in the same macro library as the median macro at http://www.oseda.missouri.edu/mscdc/sasmacro/agg.sas.

Creating a Correlation List for Wire Centers

We have discussed the tools and processes that can be used to allocate and aggregate summary data based on using a correlation list that shows the relationship between source geographic units (for which we *have* data) and target units (for which we *want* data). And we have discussed the importance of the "090" (split block group) combination of geographic levels because it is the smallest unit for which we can get 1990 sample data summaries. For this to be relevant to our problem -- getting 1990 STF3 tables at the telephone Wire Center geographic level -- we would need to have a Wire Center variable in the MABLE database. Then we could use the Geocorr engine to generate the correlation lists - one per state - needed to perform the allocation process.

Using a GIS to Link Census Blocks to Wire Centers

Our real challenge was to find a fast and reliable way to link over 7 million census blocks to Wire Centers for the entire U.S. The work that would be involved in obtaining paper maps or legal descriptions of these areas and then relating those to census maps with the block level geography would be prohibitively expensive and error prone. We needed to find a shortcut.

We were able to purchase a Wire Center boundary file from On Target Mapping (now owned by Map Info Corp.) This file contained the machine-readable definition of the Wire Center areas. It was obtained in a format that could be loaded into the *Arcview* desktop geographic information system (GIS) software. We loaded and processed one file per state. Once the WC boundary file for a state was loaded we loaded a data points layer from a file created using MABLE; this file contained the census block codes and the latitude, longitude coordinates of the block's internal point ("centroid".) Arcview was then used to perform a standard point-in-polygon operation that determined which of the Wire Center polygons "contained" (i.e. had located within its boundary) each census block centroid. Arcview appended the polygon identifier (Wire Center code) of the enclosing polygon for every block centroid point. We now (following a simple export operation) had a file with each of the blocks for the state and a corresponding WC code.

Note that in that last sentence we said "*a*" WC code rather than "*the*" WC code. Unfortunately, we discovered that in some areas there were WC polygons that were not (as with most GIS polygon coverages) mutually exclusive (i.e. non-overlapping.)

Water Clipping Glitch

We encountered another slight problem with the algorithm described above for assigning each block to a Wire Center. On Target did a special editing operation on their boundary files to make the boundaries of coastal WC's follow the shore line rather than go out to the actual boundary of the polygons. ("Coastal" here includes entities along the shore of any major body of water, not just oceans.) This is generally desirable for the purposes of mapping data, since it makes the shape of areas look more like what you would expect. But it creates problems for our point-in-polygon schemes when the centroid coordinates for blocks which were located on islands off the shore fail to be detected as being inside any Wire Center polygon. At first we thought we could get around this problem by just manually reviewing each case and assigning each island to the nearest on-shore WC, but this proved impossible in most cases. The islands are frequently adjacent to shorelines with multiple WC's and there is no way to tell to which of these the island belongs.

Post Processing to Get the Correlation List

We exported the information about block/Wire Center correspondences out of Arcview and back into SAS (the software package we used to do all our general data allocation processing.) We did not actually add Wire Centers to the MABLE database and use geocorr to generate the needed split-block group to Wire Center correlation list. Instead, we wrote a SAS program that performed the same processing in a more efficient manner and ran this program for each of the 50 states. The key was attaching the WC codes to the blocks. Once that was done, manipulating the correlation list to work with any combination of census geocodes was a solved problem (this is what *geocorr* does).

Generating the Wire Center Data

We have already discussed the processing involved in combining a correlation list with a source-geocode level data file to create a target-geocode level data file. Once we built our split-BG to Wire Center correlation list it was routine to combine this with the appropriate 1990 STF3 summary level 090 data summaries in order to create the Wire Center summaries. Of course, it helped that we already had all the 090-level data for the entire US available and in SAS data sets. This was the result of the work done by Hendrik Meij at CIESIN/SEDAC (with a little help from the author) in creating the Archive of Census Related Products at CIESIN. As part of that census data archive, they created complete directories of summary level 090 data for the entire country as compressed SAS transport files that could be FTP'ed from their site.

Reliability of the Data

Questions have been raised concerning the appropriateness of allocating the census summary data to some very small Wire Centers. The rule of thumb when doing such allocations is that you should not allocate relatively large units to relatively small ones. In other words, we want to avoid small allocation factors in our correlation lists. To this end, we chose to use the smallest possible units ("split block groups") as our source geography. In order to measure the reliability of the method for each Wire Center summary observation we create, we added code to the allocation process that calculates a population-weighted average of the "allocation factors" used to create the observation. For example, suppose we had a very small Wire Center that was created by allocating 60% of one split block group with a population of 300 people, 40% of another split BG with 200 people and all (100%) of a third block group with only 150 people living in it. In this case the weighted allocation factor statistic would be:

$$(.60*300 + .4*200 + 1.0*150) / (300 + 200 + 150) = .63 .$$

There would be 410 people in the Wire Center (the numerator in the formula) and the data allocated to those 410 people was based on data for the three split BG's with a total of 650 people. Of those 650 people, 410 are really in the Wire Center and 240 are not. The best we can do here is disaggregate data for the two split BG's not entirely in the Wire Center and come up with this estimate. It will be a "bad" estimate if the characteristics of the 300 people in the first split BG and/or the 200 people in the second split BG are not rather uniformly distributed across the portions that are inside/outside the Wire Center. Is this likely to be the case? Remember that the split BG areas, by definition, cannot be partly urban or rural, cannot be partly in and partly out of a town (because urban/rural and place are part of the split BG hierarchy.) We have two things to consider when trying to judge reliability here:

- If the value of the weighted allocation factor is at or near 1 we know we do NOT have a problem.
- Even if the weighted allocation factor is small (say, less than .5) it does not mean that the data is necessarily "bad". It would depend on the homogeneity of the split BG's that were being

allocated. It seems intuitively unlikely that such small geographic areas would have seriously differing characteristics that followed Wire Center boundaries.

Average Allocation Factor Analysis for Illinois

To determine the reliability of our allocation to Wire Centers from split blocks groups, we did a quick analysis of the average allocation factor ("*avgafact*") measure for the Illinois file. (We chose Illinois because it is large enough, has a good mix of urban and rural and tends to be a "normal" state.) There were 991 WC's in Illinois. The avgafact value was less than .5 for only two of these. The value was 1.0 (no disaggregation at all) in 466 (47%) of the cases and was greater than .9 in 449 more (45.3%). The median allocation factor was .991 (meaning half the WC's had an avgafact of .991 or more) and the average allocation factor was .981. Fewer than 1% of the WC's had avgafacts less than .88. We also looked at these same statistics broken down by WC's that had at least 50% of their populations classified as rural (per the 1990 census table and census definition). We were looking for possible bias for the more rural areas. We found only trivial differences. Of the 991 WC's, 678 qualified as rural. Within this group the avgafact statistic had a population-weighted mean of .971 (vs. .982 for urban WC's). 90% of the rural WC's had an avgafact of .89 or more. Our conclusion was that we definitely did *not* have a problem with our allocation method because the split block group geography was small enough to avoid significant disaggregation when going to the Wire Center level. (As a further check, we ran the same analysis for Pennsylvania and got very similar results.)

Allocating Data to Rural Service Areas

Overview

Once we had our data aggregated to Wire Centers, our next goal was to take these tabulations and aggregate them to areas serviced by rural telephone companies. For this we needed to be able to get the geographic relationship between the Wire Centers and the areas served by the rural companies. We had a number of data sources that would allow us to determine these relationships. The common key was the "npa-nxx" codes -- or what we called the "telephone exchange" (3-digit dialing prefix within area code within state.) From our data vendor (On Target Mapping) we had a file that provided a set of telephone exchanges that corresponded to each wire center. From NECA we obtained a set of files that defined the specific telephone companies that were designated as having rural service areas, as well as files that could be used to link all telephone exchanges with all companies. Combining all these sources would permit us to determine the approximate proportion of each Wire Center that was associated with one of the rural service providers. We would use this correspondence to do another data allocation to apportion Wire Center data to what we called the "Rural Service Area" (RSA) portions of each state and the U.S.

Creating a Wire Center to RSA Correspondence

The creation of the "correlation list" linking Wire Centers to the portion of each state serviced by a rural telephone company was rather messy. It involved doing the following set of linkages:

1. Using a file provided by On Target Mapping we were able to relate telex codes (i.e. 3-digit dialing prefixes within area codes within states) to Wire Centers.
2. Using "Tariff 4" files provided by NECA we were able to relate portions of telex areas to specific service providers (telephone companies.) We were somewhat surprised to learn that telex's could be linked to more than one telephone company but that turned out to be the case. The file

that provided the linkage specified a range of 4-digit dialing codes for each of the telex-company combinations. We used this information to estimate the portion of each telex that was serviced by a company. The large majority of exchanges (71,855 out of 72,891 or 98.6%) were serviced by a single company.

3. Using another file from NECA that identified the specific rural telephone companies whose service areas were to be the definition of "Rural Service Area" for our purposes we were able to tag each telex/company combination as being either in or not in one of these RSA's. This was a simple matter of merging the list from above (telex to telco) with our list of specified rural telco's and recording whether or not we had a match.

At this point we had a list of telex codes that went with Wire Centers and another list that told us which portion of those telex's were associated with RSA's. We were able to combine these two sets of information to determine what portion of each Wire Center was associated with RSA's. Because we wanted to focus on areas that were entirely or at least almost entirely rural, we decided to designate Wire Centers as serving a rural area if we estimated that at least 70% of the phones within their associated telex codes were linked to RSA's. We created a Wire Center to "Rural Tel Co" correspondence file that had no allocation factors. A Wire Center was to be classified as entirely rural or not based on this 70% criteria.

Creation of the RSA Portion Data

We were now ready to use the data that we had already allocated to Wire Centers and the correlation list that defined Wire Centers as being in or out of the Rural Service Area portion of a state. A simple merge and aggregate (no allocation factors this time) process was done to get data summarized for the RSA portion of states, i.e. the portion that was served by one of the rural telephone companies. A summary report was generated and can be viewed at http://www.oseda.missouri.edu/telexdata/agg 2rural2.summary.report. It shows, among other things, that:

- 7% of the US population was contained in Wire Centers that we were able to classify as being in Rural Telephone Service Areas.
- 92.3% were in areas classified as being non-Rural Telephone Service Areas.
- 1% (283,919 persons) were not assigned to either category because they could not be assigned to a Wire Center. We refer to this category as our "island population" since it results from the inability to assign off-shore census blocks to Wire Centers.
- .5% (about 1.3 million persons) were not assigned to either category because we were unable to link the telephone exchange codes from our vendor data (relating them to WC's) and those on the Tariff 4 data.